ASSEMBLING A HIGH-REDSHIFT GALACTIC PROTOCLUSTER: AN IMAGING STUDY OF MRC1138-262

by

JONAH ROSE

A THESIS

Presented to the Department of Physics
and the Robert D. Clark Honors College
in partial fulfillment of the requirements for the degree of
Bachelor of Science

June 2019

An Abstract of the Thesis of

Jonah Rose for the degree of Bachelor of Arts
in the Department of Physics to be taken June 2019

Title: Assembling A High-Redshift Galactic Protocluster: An Imaging Study of MRC1138-262

Approved: _____

Dr. Scott Fisher

In this thesis we present multiband optical and near infrared imaging observations using FourStar and DECam of an extended field around MRC1138-262, the central galaxy in a forming galactic cluster at a redshift of 2.16. The central part of this field contains a spectroscopically confirmed overdensity of galaxies. We use multiband photometry to identify two new potential cluster members and 16 new field members at redshift near 2.16. Using these galaxies and previously spectroscopically analyzed galaxies, we begin to compare the properties of the cluster galaxies to those in the field. We confirm a significant overdensity of galaxies in the core of the cluster and an overdensity of Ly-alpha emitters. We do not find a significant change in the distribution of red galaxies in the cluster, suggesting a weaker link than previously thought between cluster galaxies and H-alpha emissions.

Acknowledgements

I would like to thank Dr. Ricardo Demarco for supplying the data necessary to complete this project and the advice he has provided throughout the process. I would like to thank Dr. Inger Jørgensen for her knowledge, wisdom and guidance that has allowed this project to flourish, without which I would be lost. I would also like to thank my Thesis Committee that includes Dr. Scott Fisher for his support, enthusiasm and advice that have helped me to gain footing in this intimidating field; Professor James Imamura, who served as my Second Reader; and Professor Mark Carey, who served as my Clark Honors College Representative. I would like to thank Dr. Mischa Schirmer, Dr. Masayuki Tanaka, and Maggie Thompson for their information and analysis which allowed this project progress to its current state. Finally, I would like to thank my friends and family for their continued help and support.

Table of Contents

List of Figures

List of Tables

Introduction

Within the broad field of extragalactic astronomy, galaxy protocluster study has recently developed due to advancements in image processing. With the construction of telescopes with 8-10 meter apertures and the rapid advancement of digital imaging sensor technology, astronomers are routinely making observations that were impossible 10 years ago. The new technology has opened new facets of observational astronomy and allowed us to probe previously inaccessible parts of the observable universe. In this paper, we utilize these advancements to better understand the formation and evolution of a galaxy protocluster surrounding MRC1138-262.

Galaxy protoclusters, a grouping of galaxies that will become gravitationally bound, are difficult to locate due to their rareness and dimness. Since these clusters are so far away from Earth, the closest having a redshift (z) of 1.5 which corresponds to being ~8 billion lightyears away from our solar system, even the brightest cluster galaxy (BCG) can be too faint to appear in imaging surveys. Even when protoclusters are found, astronomers must devote significant telescope time to detect the fainter cluster galaxies in the vicinity of the BCG. The detection issue is more pronounced when performing spectroscopy, the study of how an object emits light over different wavelengths. However, astronomical technology has developed enough to obtain images and spectra of these clusters if we know their location in the sky.

There are two data types that researchers use to study protoclusters: imaging and spectroscopy. Both data types must be collected at an observatory, but each uses a different type of instrument to collect light. Imaging data is collected by placing a light-sensitive detector, like a charged coupling device (CCD), at the focal plane of a

telescope optical system. After leaving the telescope, the light passes through a filter which only lets through certain wavelengths of light, and then hits the CCD camera to produce a monochrome image. Spectroscopic data is taken in a similar manner, except the light is dispersed by a grating before it enters the CCD, revealing how the intensity changes with wavelength. Spectra provide much more detailed information about scientific objects which can be used to determine properties that cannot be understood from imaging data alone, such as dynamical mass and metallicity (Padmanabhan et al. 2003). Although spectra can provide much more information about an object, it requires much more telescope time than imaging data, which is difficult to obtain at major observatories. The data used in this project only includes imaging data; however, the results will help future efforts obtain spectra of galaxies in this protocluster by identifying potential targets for spectroscopic observations.

To facilitate spectroscopic analysis of the infalling region around MRC1138-262, we use broad-band and medium-band photometry to identify likely protocluster members. Photometry allows astronomers to investigate many more objects than spectroscopy because photometry requires much less observing time per unit area. Due to spectroscopy's requirements, professional observatories with 8-meter class telescopes are unlikely to devote the time required to gather spectroscopic data on every galaxy present in a 11x11 arcminute field around the BCG. This limitation is why the fields in the previous studies have been spatially limited to the area surrounding the protocluster. To make spectroscopy of this protocluster more feasible, my work will reduce the number of objects that must be spectroscopically imaged to only ones that have a high likelihood of being spatially associated with the BCG. We predict that this work will

reduce the number of potential sources from thousands to tens of galaxies thus greatly reducing the time necessary to collect the spectroscopic data necessary to understand how the protocluster around MRC1138-262 formed.

An important and not well understood aspect of this protocluster is the infalling region. The infalling region is the space around the BCG where objects are gravitationally influenced by the cluster, but they have not yet interacted with other cluster members due to their distance from the BCG. Galaxies in this region will eventually become part of the cluster as the gravitational force from the cluster pulls galaxies in this region toward the center of the protocluster. Studying the properties of galaxies in this region can help astronomers understand how the environment around protoclusters will affect the cluster's evolution.

MRC1138-262, known as the Spiderweb Galaxy, is the BCG for one of these protoclusters. It is located at z=2.16 (corresponding to a lookback time of ~11 billion years) and is one of the brightest and largest protoclusters known to exist at this extreme distance from our galaxy. Because of its brightness, this protocluster has been heavily studied (e.g. Pentericci et al. 1997) to understand the characteristics of galaxy cluster formation. This paper will extend this work by attempting to answer the question, how many protocluster members and infalling-region members can be detected in an 11x11 arcminute region around MRC1138-262? We estimate that ~30 new sources will be detected in the infalling region of MRC1138-262.

The general question astronomers have been trying to answer about the Spiderweb Cluster, and protoclusters in general, is, how does the evolution of galaxies in clusters differ from those not in a cluster (i.e. so called "field galaxies")? To address

this question, astronomers have obtained spectra as well as imaging data with multiple filters of galaxies in the protocluster. So far, astronomers have used the collected data to understand characteristics of the cluster system, such as satellite morphology and color. However, the data available for the Spiderweb protocluster are not sufficient for astronomers to understand how the Spiderweb cluster formed or how it will evolve. The results of this paper will complement existing data to improve our understanding of how the Spiderweb protocluster formed as we will have a better understanding of the environment around the protocluster itself.

To measure galactic characteristics, astronomers use spectroscopy to understand the galaxies' chemical composition and imaging to understand their morphology. Kurk et al. 2003, Miley et al. 2006 and Hatch et al. 2009 imaged a 5 x 5 arcminute field surrounding the Spiderweb Galaxy with the goal of detecting galaxies that are Lyman Alpha emitters (LAEs) and Hydrogen Alpha emitters (HAEs). If LAEs are found, then the galaxies are likely very young and undergoing their first burst of star formation. If HAEs are found, the galaxies are most likely older and in subsequent star forming periods. Kurk et al. 2003 found both kinds of emitters around the protocluster, raising the question to how this cluster formed. Miley et al. 2006 also uses Lyman Alpha detections to compare LAEs to other galaxies in the cluster that are not emitting Lyman Alpha. Miley's paper shows that linear galaxies, whose morphologies are described as "chain" or "tadpole", are most often LAE and were more common in this protocluster than would be expected to find in the field. The overdensity of LAEs around the BCG indicates a link between LAE and galaxy cluster formation but requires a larger sample to fully understand.

Another prominent feature of the protocluster around the Spiderweb that has been well-studied is its red sequence (Kurk et al. 2004a, Zirm et al. 2008, Kodama et al. 2007). For a protocluster, the red sequence is the group of galaxies present in the protocluster that are redder than what would be expected if they were found in the field. Kurk et at. 2003 obtained spectroscopy of galaxies around the Spiderweb to identify LAEs, HAEs, X-ray emitters and extremely red objects (EROs). The paper found an overdensity of EROs around the Spiderweb, indicating the presence of a strong red sequence. There are two likely reasons why an object can appear red in a cluster, either they are reddened by the intracluster medium (ICM) or have had their star formation stopped by the BCG. However, since the spectroscopic sample size was so small, we do not have spectra for enough of these galaxies to determine why these objects are red. The two different reasons for a reddened galaxy can indicate different reasons for how this cluster formed, either through merger of two systems or a single system that has had some if its star formation stopped in the past. Kurk et al. 2004b theorizes that the latter scenario is more likely because the crossing time of the galaxies is larger than would allow the protocluster to enter a relaxed state. Determining which process is reddening these galaxies can help determine the protocluster's formation process by determining how satellite galaxies have interacted with the BCG.

In previous studies of the Spiderweb cluster, imaging data from the optical to near infrared have been collected for the central region of the cluster, at most 3.2 x 2.5 arcminutes around the BCG (Zirm et al. 2012). While these observations have provided details about the detected galaxies, they focused on the core of the protocluster. The small region was chosen because there has not been a systematic study to determine

which galaxies further away from the core are part of the protocluster. This paper will extend the information on galaxies in the outer region and in the infalling region by determining the distances between the galaxies in the field and the BCG. This process will be done by using imaging data from 7 different filters to obtain photometric redshifts of all galaxies in a 11'x11' area around the Spiderweb Galaxy.

Methods

Data Collection

Data collection was performed by our collaborator Dr. Ricardo Demarco of the Universidad de Concepción, Chile, over two nights in February 2013. Using the four-meter telescope at Cerro Tololo Inter-American Observatory (CTIO) and the 6.5 meter telescope at Las Campanas Observatory (LCO), Dr. Demarco obtained imaging data on MRC1138-262 in seven filters, five broad-band and two medium-band, along with images of calibration stars. The wavelength coverage of the images ranges from the optical to the near-infrared. The system throughput, defined as the percentage of light that passes through the optical train, is shown in **Figure 1**.

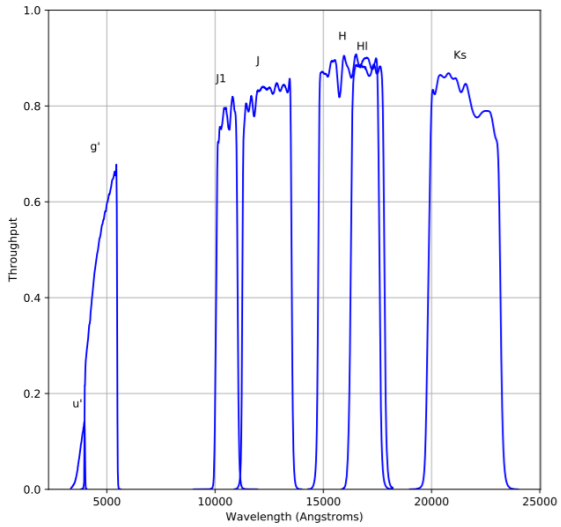


Figure 1: System Throughput

System throughputs as a function of wavelength for each of the seven filters used to collect data. The name of the filter appears above its curve. The system throughput describes how much light from an astronomical object will be detected by the camera after being attenuated by the atmosphere, filter and the instruments' optical systems.

The optical images were obtained on the Blanco 4m telescope using the Dark Energy Camera for Surveys (DECam) at Cerro Tololo Inter-American Observatory, Chile (Diel et al. 2012). The near-infrared (NIR) images were collected on the Magellan Baade 6.5m telescope using the FourStar camera at Las Campanas Observatory, Chile (Persson et al. 2013). Information on these images can be found in **Table 1**.

Filter	Number of Images	Seeing (arcseconds)	Total Exposure Time (s)	Pixel Scale (arcseconds/pixel)
u’	6	1.51	5762	0.2636
g’	6	1.46	2030	0.2636
J1	17	0.41	698	0.159
J	30	0.41	1222	0.159
H	472	0.56	4112	0.159
Hl	312	0.43	3500	0.159
Ks	216	0.40	3130	0.159

Table 1: Image Characteristics

Column 1: The name of each filter. Column 2: The number of images that were taken in each filter. There are fewer images in the optical (u’ and g’) filters because each image has a longer exposure than in the NIR bands. Column 3: The average width of a point source (star) for each filter. Column 4: The total exposure time for all images taken in that filter. Column 5: The plate scale of each instrument.

Data Files

The data for this project consist of approximately 75 GB of images and auxiliary files. The image files from both instruments are in the FITS format (Wells 1979), which is standard in professional astronomy. FourStar consists of four detectors, a layout is shown in **Figure 2**. Each detector has 2,000 x 2,000 pixels and the camera’s overall field of view is 10.8 x 10.8 arcminutes. Between each detector is an 18 arcsecond gap which is not imaged.

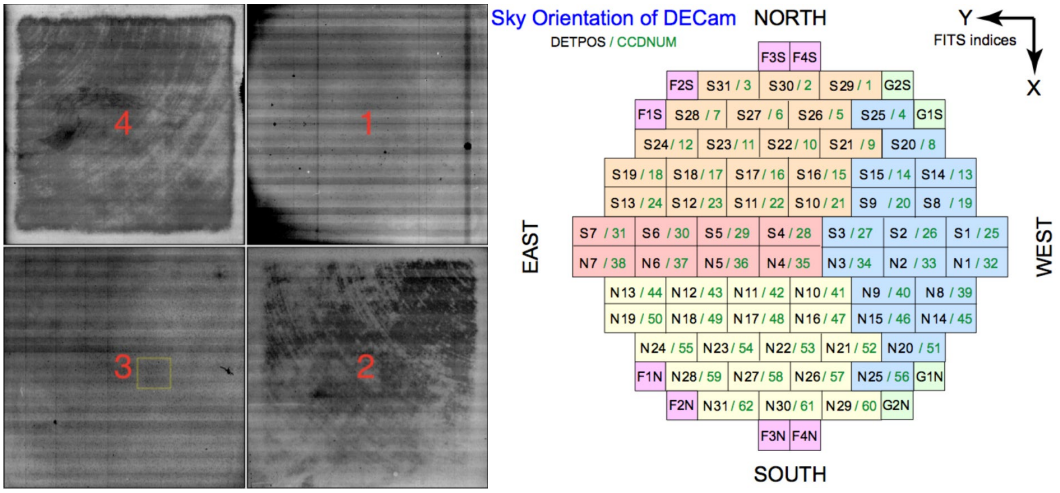


Figure 2: Detector Layout

Spatial layout of detectors on the FourStar (left) and DECam (right) cameras. Each detector acts as an individual camera that detects the light that travels through the telescope system. Astronomers use the numbers on each detector to index FITS files. [Image credit: http://instrumentation.obs.carnegiescience.edu (left) & http://www.ctio.noao.edu (right)]

DECam is one of the largest imaging arrays on any telescope and consists of 62 detectors used for imaging science objects, 8 detectors used to maintain focus, and 4 detectors used to help guide the telescope. A layout of these detectors is shown in **Figure 2**. The focus and guiding detectors, labeled with an "F" and "G" in **Figure 2,** are not useful during data collection and therefore their data are not included in the data files. Two of the science detectors were not functional at the time these data were collected. Each science detector is 2k x 4k pixels, the gap between rows of detectors is 53 arcseconds and the gap between columns of detectors is 41 arcseconds. The entire instrument's field of view is a circle with a radius of 1.1 degrees.

Along with the imaging data, Dr. Demarco provided a catalog containing photometric data that was created by Dr. Masayuki Tanaka (Tanaka et al. 2013). This catalog contains imaging data in nine filters in a 5 x 5 square arcminute region centered

on the Spiderweb galaxy. Five of the filters are also included in the Demarco imaging data. This data will be used to compare to the reduced and calibrated images produced in the next section.

Image Reduction and Calibration

Overview

To analyze the characteristics and properties of the data, we first remove systematic and random noise inherent to astronomical imaging data using five types of images. The five types of images are: biases, darks, flats, science, and standard star. Each of these images are described in the next subsection. Bias, dark, and flat images are used to reduce systematic noise in the science and standard star images. Standard star images are used to calibrate the science images. Often, science images combined to create a master science image which astronomers analyze to answer scientific questions.

The image reduction process is done separately for each filter, the result of each reduction will result in one master image that has the best signal to noise (S/N) possible. To perform this process for this data, I used the software program THELI (Schirmer et al. 2005; Schirmer 2013). THELI is an analysis tool used by astronomers which has the camera properties, reduction scripts, and stellar catalogs built into a pipeline to make the data reduction process simpler and allow astronomers to fine-tune the reduction process for their data.

The overall processes of THELI used to reduce this data will be described in the following sections. The process begins with the raw science and calibration images. The first step is to combine the similar calibration images to create master calibration

images. The second step is to remove background noise from the science images. At this point, all systematic noise has been removed from the science images. The third step is to create weight maps for the science images. Weight images tell astronomers how much information is contained in each pixel of an image. The fourth step is to correct for astrometric (spatial) distortions. The fifth step is to set the average value of the background to zero then combine the science images and the weight images. At this point, the master science image and weight map have been created. The final step is to use the standard star data to measure zeropoints for each filter. The zeropoints are a way to calibrate the master science images to a standard system so that results derived from these images can be compared to other publications.

Image Type Description

When collecting data, there are five types of images that must be collected to produce research-grade images. These five types are bias, darks, flats, science, and standards. Bias, darks, and flats, henceforth referred to as error correcting images (ECIs), are used to diminish different forms of background noise that are inherent to the detectors used in these observations from the science and standard images. Multiple images are taken for each kind of ECI, usually an odd number between 3 and 11, then median combined to create a master image that reduces random noise from the detector.

Bias images are taken with the shutter of the camera closed and exposed for zero seconds. These images measure the effect of the initial electrical flow to the camera on the detected image. An example of a bias image is shown in **Figure 3**. Dark images are like bias images in that they are taken with the camera shutter closed, but their exposure time should be the same as the science and standard images. These images detect both

the initial electrical flow to the camera and the continual energy flow to keep the camera on. An example of a dark image is shown in **Figure 3**.

Flat images are taken with the shutter open and looking at an object with uniform illumination. The uniform illumination is obtained either by looking at a portion of the sky with no objects in it during twilight, or at a flat white screen located somewhere in the telescope dome. These methods are called sky flats and dome flats respectively. Sky flats are taken with telescopes that have a small field of view because it is possible to find sky patches that are flat over their small window. Large telescopes use dome flats because it is difficult to find a flat patch of the sky that does not contain any bright sources that would appear in the image. The purpose of flat images is to reduce the effects of dust particles or other obstacles in the optical path of the telescope and camera system which block some light reaching the camera. Flat images also correct for sensitivity differences across the face of the detector. An example of a flat image is shown in **Figure 3**.

Science images are images of the sky used to answer a scientific question. Images are taken with different filters, exposure times, and fields of view. The exact combination of filters, exposure times, and fields will depend on the design of the observing project. At the level of this data, the design of the project is a direct consequence of the scientific question being answered. Photometric filters can span from radio waves, with the longest wavelength light, to X-rays, with the shortest wavelength. The wavelength bands astronomers use allow them to understand different characteristics of objects. The exposure time determines the limiting magnitude, or how bright an object must be to be guaranteed to be detected. Bright objects, like stars,

require less time to image than faint objects, like far away galaxies. The field of view represents the area of the sky where the telescope is looking. For larger objects, like a cluster of stars or galaxies, a wider field of view is necessary to have all the objects in the field. Smaller objects, like single stars, require a much smaller field of view. An example of a raw science image is shown in **Figure 3**.

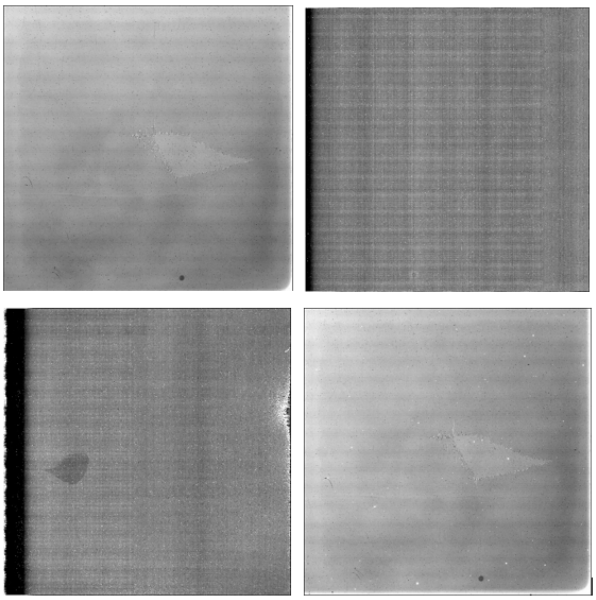


Figure 3: Image Types

Examples of each image type used during data reduction. A Flat Field is shown top left which is used to correct for multiplicative errors due to detector efficiency. A Bias is shown top right which is used to correct for noise from the initial electrical charge sent to the camera. A Dark is shown bottom left which is used to correct emission from the continual usage of the camera. A science image is shown bottom right from which is what the ECIs remove their corresponding noise. All four images were taken with FourStar.

Standard stars are images of specific stars with a known brightness in each filter used to create science images. The purpose of these images is to measure how much light the camera detects from each of these stars and use that information to understand

the true brightness of the objects imaged in the science images. It is important for astronomers to choose many standard stars of different colors to understand if the camera sensitivity depends on the color of the light it detects.

Step 1: Calibration

The first step in this process is to create master images for each of the calibration types. Since biases, darks, and flats do not depend on where the telescope is pointing, this process is straight forward. A master calibration file is made by looking at an individual pixel in each image, taking the median value of the list of pixel values and giving that pixel in the master image the median value. This process is performed for every pixel in the image to create one master image that has the same dimensions as the individual images. Taking the median of each pixel will remove outliers and choose a value that better represents the true brightness at that position, thus reducing the random noise in each master calibration file.

The systematic noise captured by the calibration files is then removed from the science images using **equation (1)**. Dividing by the normalized flat also normalizes the gain across all detectors. This is not important for the data reduction process but will be helpful when the catalogs are created.

$$Corrected\ Image = \frac{(Raw\ Image) - Bias - (Dark - Bias)}{\mathrm{Norm(Flat)} - (Dark - Bias)} \quad (1)$$

Step 2: Background

Background correction removes light from the detector that was collected from the sky during observations and removes any background structure present in the image. The first step in this process is to determine which pixels in the image are part of the

background and which are part of an object. This is done by measuring the variance in pixel values. To first order, any set of pixels that has at least 3 connected pixels that are at least 3 times the standard deviation of the pixels is determined to be an object. These pixels, and the surrounding pixels, are removed (masked) from the image for the purpose of determining the background. This process is then preformed iteratively to make sure that the faint objects are also masked.

Once the objects have been masked in each individual image, the remaining pixels constitute the background of that image. The median of the background images creates a master background image for each detector. Finally, the background is removed from the images from which it was measured.

Removing the background from the images can be done in two ways. The first way is to subtract the master background image from the science images. This will remove additive effects, such as airglow or scattered light. The other way is to divide it. Dividing the background from the science image is necessary if the flat field was not made correctly. A bad flat-field is more likely to occur in wide-field cameras because it is harder to get even illumination across the entire field of view.

For our data, the background was subtracted from both the near-infrared and optical data. This is standard for the near-infrared but was chosen for the optical data as well because the only the area that overlaps with the near-infrared images will be analyzed. Therefore, it is not necessary to correct for uneven illumination across the entire field of view.

Step 3: Weighting

The purpose of this process is to create weight images for the individual science images. A weight image describes how much information each pixel contains and is helpful for determining error bars and normalization. The exposure time and quality of the detector are two factors that determine how much information is stored in a pixel. The first step in this process is to create a characteristic weight image, called a global weight image, from the master flat image created in Step 1. The global weight image measures how much information is collected by each pixel on a particular detector. This is done by setting the value of each pixel in the weight map to the corresponding pixel value in the normalized flat.

Once the global weight images have been created, they are modified by individual images to create weight images that will be used during coaddition, Step 5. During this step, a trained neural network is used to detect cosmic rays, depicted in **Figure 4**, in the individual image and set the weight of these pixels to zero. Since cores of bright stars consist of many connected bright pixels, the neural network can mask them as well. We did not avoid this effect because we do not use these bright stars in our analysis.

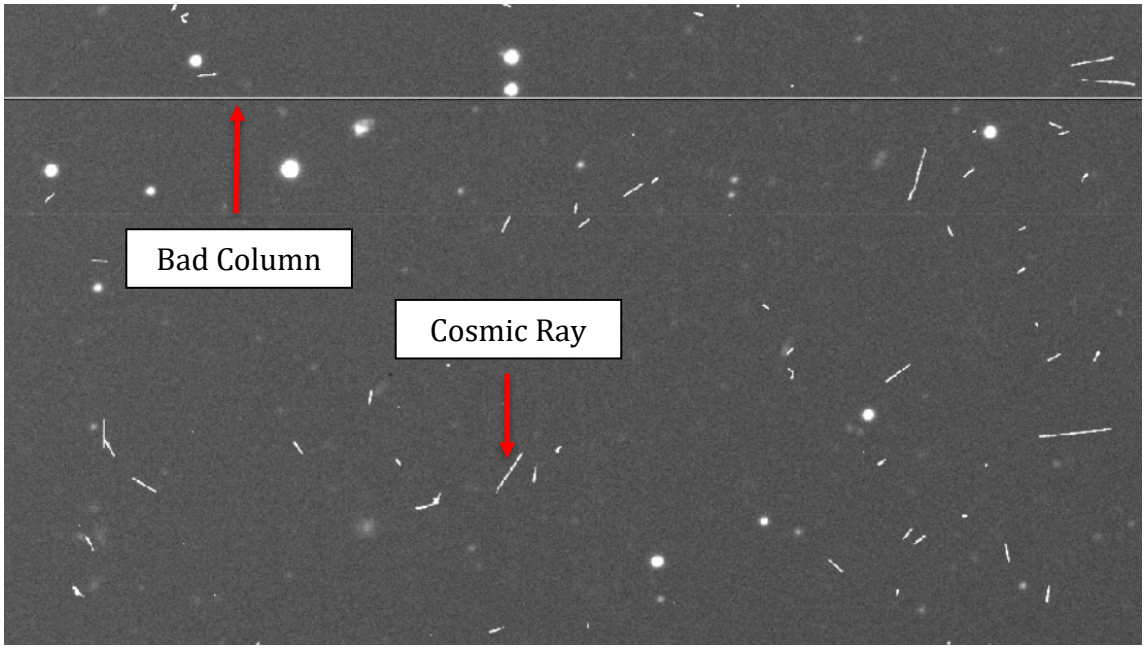


Figure 4: Cosmic Rays

Science image from DECam before it is correct for cosmic rays and bad pixels. The cosmic rays are atomic nuclei that impact the detector and appear as bright streaks across the image. Bad pixels can either appear as single or lines of bright or dark pixels and are pixels that do not detect photons at the same rate as the rest of the detector.

Step 4: Astrometry

The purpose of the astrometry step in THELI is to identify where on the sky the telescope was pointing for each image and to correct for distortions in the image. This can be done by comparing the bright objects in each image to the same bright objects found in large sky surveys with well-known coordinates in the sky. We used the GAIA (Gilmore et al. 2012) survey to match the positions of our stars because it has imaged the MRC1138-262 field in both optical and near infrared wavelengths. Once we determine the location of our image on the sky, we correct for spatial distortions in the image. These distortions arise from taking a flat camera and trying to image a curved surface (the sky). The distortions calculated for DECam and FourStar are shown in

Figure 5. After correcting for this distortion, seeing limits the positional accuracy of the objects rather than system errors.

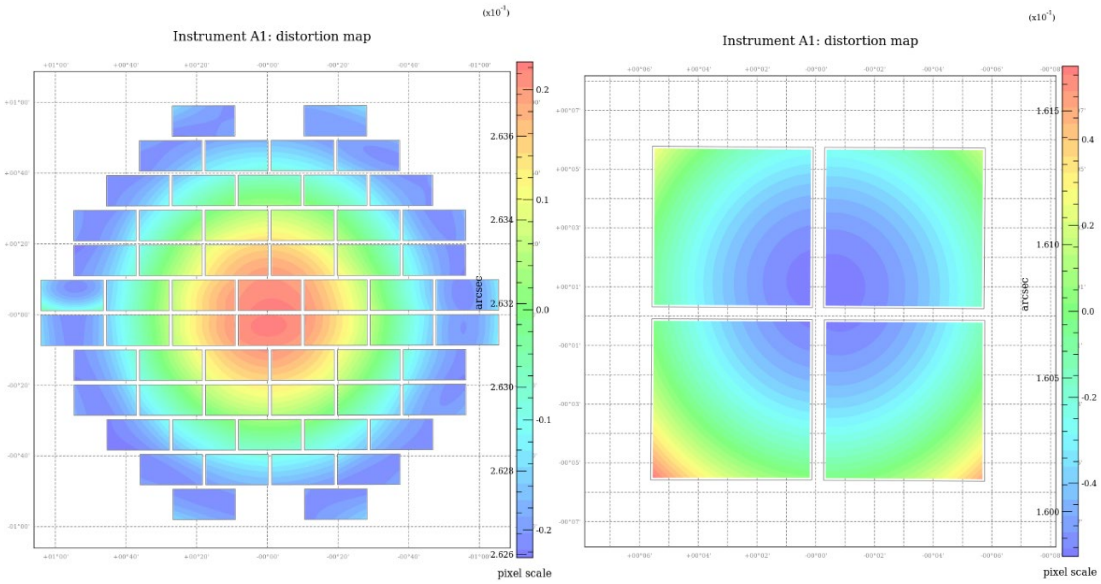


Figure 5: Astrometric Distortions

3rd Order astrometric (spatial) distortion maps for DECam (left) and FourStar (right). The distortion arises from projecting a round surface (the sky) onto a flat surface (the detector). The distortions are measured in tenths of arcseconds. The distortions were corrected to the pixel scale values found in **Table 1.** The unusual distortion found in the left-most detector on DECam results from a broken amplifier. This detector was not used in the data analysis.

Step 5: Coaddition

The desired outcome of coaddition is to take the many cleaned individual images and combine them into one master image for each filter. First, any residual light from the sky is subtracted from each image by setting the median background value to zero. The process of removing this light is similar to removing the background, but the background is interpolated across the now-masked objects. Then, an average is taken of the now-complete sky image and this average is subtracted from the science image.

Finally, all science images for each filter are aligned based on their sky position and each pixel is averaged with all other pixels that overlap with the same portion of the sky as the original pixel. The image created by all the averaged pixel values is the final master science image for that filter.

At this point, the errors in the imaging data have been fully reduced and all that is left is calibration. **Figure 6** shows a three-color image made from all seven master images created after this step. In the image, the Spiderweb Galaxy is identified, and a length scale is placed in the lower right.

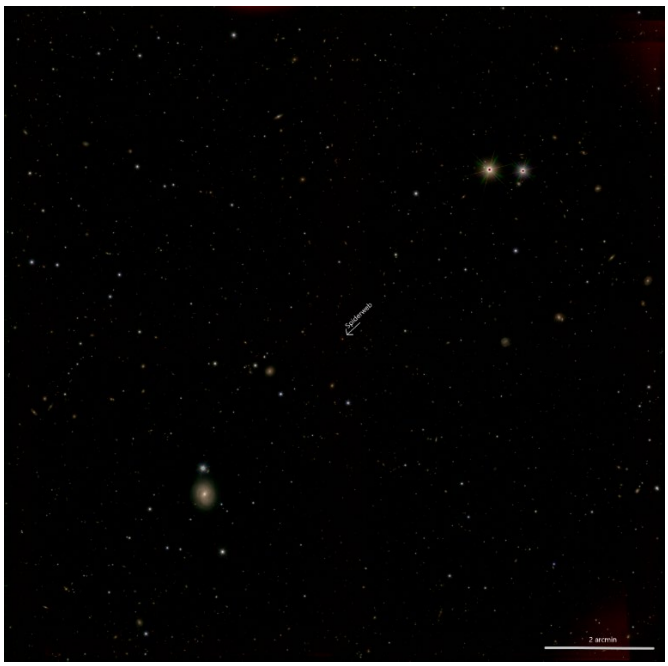


Figure 6: Reduced Three-Color Image

A Three-color stacked image of the field overlapped by DECam and FourStar made using the seven master science images created from the data reduction process. A combined u' and g' image was used as the blue color. A combined J1, J, H, and Hl image was used as the green color. And the Ks image was used as the red color. The red area in the bottom-right corner is due to a transient object and appears stronger in this image due to the coloring process.

Standard Stars

Once a master science image has been created, it is calibrated to a widely-used photometric system called the 'AB system' (Oke 1964) so it can be compared to results in other publications. This calibration is done by determining a zeropoint from the standard stars. To determine the zeropoint, the standard star images must undergo the same data reduction as the science images, but only steps one through four. Then, the total flux from the standard star is measured and related to its known calibrated value.

The flux from an object can be measured by summing the amount of light (counts per second) for each pixel that contains signal from the object. Before the measured flux can be compared to a known value, the derived atmospheric and color corrections must be applied. Atmospheric corrections occur because some light from each source is scattered as it travels through the atmosphere and does not reach the detector. The amount of light that is scattered depends on the wavelength, and thus the filter, as well as the amount of atmosphere the signal traverses. Color corrections are applied to the data due to the photon-counting nature of the detectors and because the detectors are more efficient at detecting certain wavelengths of light. Specific values for the atmospheric and color corrections can be found in **Table 2**. These corrections are applied using **equation (2)**.

$$Corrected\ Image = (Raw\ Image) + b * Color + k * X + a \qquad (2)$$

Filter	Color Constant (b)	Atmospheric Constant (k)	Galactic Constant (a)	Airmass (X)	Color
u’	0.04	0.44	0.145	1.05	(u’ – g’)
g’	0.11	0.20	0.113	1.08	(u’ – g’)
J1	0.00	0.03	0.026	1.24	N/A
J	0.00	0.10	0.024	1.24	N/A
H	0.00	0.04	0.015	1.01	N/A
Hl	0.00	0.04	0.015	1.40	N/A
Ks	0.00	0.08	0.010	1.03	N/A

Table 2: Magnitude Correction Parameters

Column 1: The name of each filter. Column 2: The magnitude correction coefficient for the detector’s color sensitivity for each filter. Column 3: The coefficient for how the light scatters in the atmosphere for each filter. Column 4: The coefficient for how the light scatters in galactic dust between the source and Earth. Column 5: The number of airmasses that the light travels through in the Earth’s atmosphere. One airmass is defined as looking at zenith (straight up). Column 6: The filters that are used to calculate the color of the object, used with the color coefficient. This column does not apply to the NIR filters because the color coefficient is zero.

Ideally, astronomers would be able to determine the correct atmospheric and color corrections for the specific night that the data is taken by measuring standard stars of different colors and different amounts of atmosphere. However, due to limitations while observing and some standard stars being saturated, the published values for the color and atmospheric corrections for each observatory site were used (Persson et al. 2002; Li et al. 2014). A summary of these values can be found in **Table 3**.

Once the fluxes have been measured and calibrated, they are compared to published values. The near-infrared standard stars were compared to the values published on the FourStar website, which were calculated synthetically using the spectra of the star in conjunction with the FourStar system throughput. The optical

standard stars were compared to the E-Regions catalog (Cousins 1973). Results of the zeropoints can be found in **Table 3**.

Filter	Star/Region	RA	DEC	Zeropoint	RMS	Number
u’	E6-A	14:45:33	-45:15:34	24.46	0.067	8
g’	E6-A	14:45:33	-45:15:34	25.58	0.034	15
J1	GD153	12:57:02	22:01:53	26.45	.011	4
J	GD153	12:57:02	22:01:53	27.03	.003	4
H	GD153	12:57:02	22:01:53	27.27	.013	8
	C26202	3:32:33	27:51:49			
Hl	GD153	12:57:02	22:01:53	26.65	0.024	4
Ks	GD153	12:57:02	22:01:53	26.85	0.017	4

Table 3: Standard Stars

Column 1: The name of each filter. Column 2: The name of the standard star (NIR) or the name of the region of standard stars (optical). Column 3: Right Ascension (celestial latitude) for each standard star measured in hh:mm:ss. Column 4: Declination (celestial longitude) for each standard star measured in hh:mm:ss. Column 5: Zeropoint for each filter in the AB system. Column 6: The root-mean-square (error) calculation for each zeropoint calculation. Column 7: The number of data points (images) for each zeropoint calculation.

Catalog Creation

Once the master science images are created and calibrated for each filter, the next step is to extract information about each object in the image. This process was done using the software program Source Extractor (Bertin & Arnouts 2010). Sources were located using the H band due to the high signal-to-noise ratio of the data. Then the parameters of all objects that were found are measured in each master image. The process of finding objects is similar to Step 2 of the data reduction process so it will not be summarized here.

The output from Source Extractor used to analyze the data are the stellaricity, auto magnitudes, and errors. Stellaricity is a measure of how likely an object is to be a

star, with zero being certainly a galaxy and one being certainly a star. Since we are only interested in galaxies for our final analysis, this parameter will be used to remove stars from the final catalog. Auto Magnitudes are Source Extractor's best estimate as to each object's total magnitude. This parameter is used during Zeropoint Comparison to provide a standard measure to compare to other researcher's findings.

Catalog Merging

After Source Extractor creates the individual catalogs for each filter, they are merged to create one catalog that contains all data that will be analyzed. Since the same objects were measured in each filter, merging the seven catalogs from Source Extractor is trivial. Once the merged catalog has been created, it must be merged with the catalog provided by Dr. Demarco. These catalogs are similar to the merged catalog but the astrometry was calibrated to a different survey so the position measurements are not the same between catalogs for the same object. To merge these catalogs, the location of every object in Dr. Demarco's catalog is compared with every object in our catalog. If there is an object that has the same coordinates within 0.5 arcseconds, the columns of this object from the two catalogs are combined and added to the final catalog. For all objects that are not matched in either catalog, their information is added to the final catalog without combining any columns from the other catalog. There were no instances for this dataset where more than one object was identified as a match.

Once the final catalog has been created, it must be calibrated for atmospheric extinction, color effects, and reddening due to galactic dust. These corrections are made for the same reasons as the corrections made to the standard star data. **Equation (2)**

gives the calculations for these calibrations and **Table 2** gives the parameters for the calculations.

Stellaricity

Note: The work summarized in this section was performed by Maggie Thompson, another member of Dr. Fisher's research group. The results and methods are included in this thesis as they are necessary to reproduce our results.

The final step before the catalog is ready to determine photometric zeropoints is to remove the stars. To distinguish stars from galaxies, we create and examine magnitude vs stellaricity plots, shown in **Figure 7**. Based on the quality of the data in the NIR filters and the results of Jørgensen et al. 2018, it was determined that all objects defined by **equation (3)** are galaxies.

$$(Total\ Ks\ Magnitude) * (Total\ H\ Magnitude) < .64 \quad (3)$$

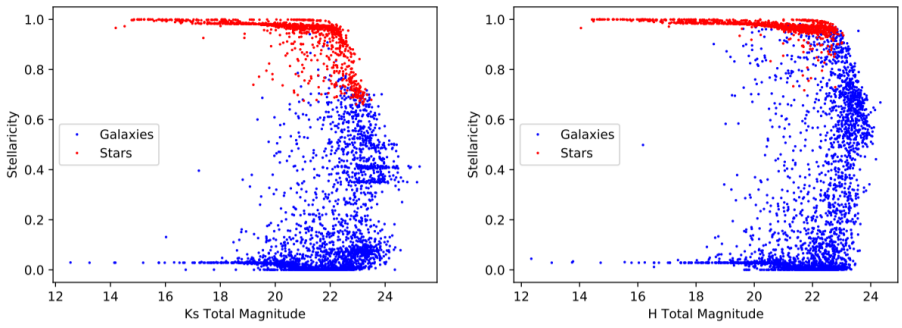


Figure 7: Stellaricity

Stellaricity plots for the Ks (left) and H (right) filters. These plots are used to determine which objects are galaxies and which are stars. Objects determined to be stars are shown in red and galaxies in blue. The horizontal groupings of galaxies with a stellaricity near 0.5 in the Ks image result from objects that are too faint to be measured accurately.

Determining Photometric Redshifts

Once the photometric catalog has been created, calibrated, and cleaned, we can use it to determine photometric redshifts for the remaining galaxies. This is done by first creating a spectral energy distribution (SED) for each object, then comparing the SED created from the catalog (observed SED) to SEDs of known galaxy types (template SED), along with sizes, and locations as additional model parameters. An example SED is shown in **Figure 8**. The template SEDs are then used to determine which galactic characteristics, such as emission lines and redshift, are best fit to the observed SED.

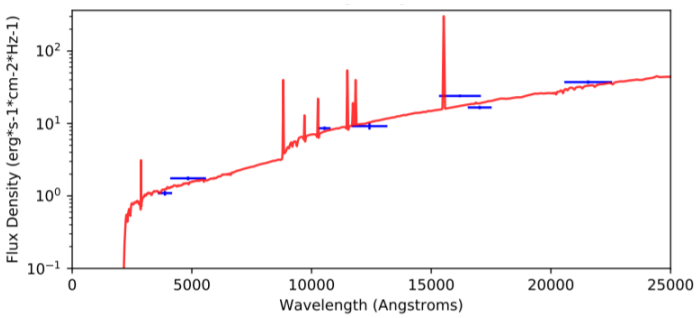


Figure 8: Example Template and Observed SED

An example of a template SED (red) and an observed SED from our data (blue). The template SED was produced by EAZY and shows the SED that best fits the data. The vertical lines that appear in the template SED are emission lines, meaning there is more flux at those specific wavelengths. On the observed SED, the x-axis errors show the width of the filter used to create the datapoint and the y-axis errors show the error in the flux measurement calculated by Source Extractor.

The software program EAZY (Brammer et al. 2008) was used to complete this process since it has a well-known database of template SEDs chosen to fit many galaxy types and built-in scripts to make and compare SEDs. EAZY takes the merged catalog as an input and outputs the observed SED and the best fit template SED. The template

SEDs, at redshift 2.16, used in this paper are shown in **Figure 9** along with the filter positions. To choose the best SED, EAZY calculates the chi-squared value for the observed SED and linear combinations of template SEDs at redshifts ranging from zero to five.

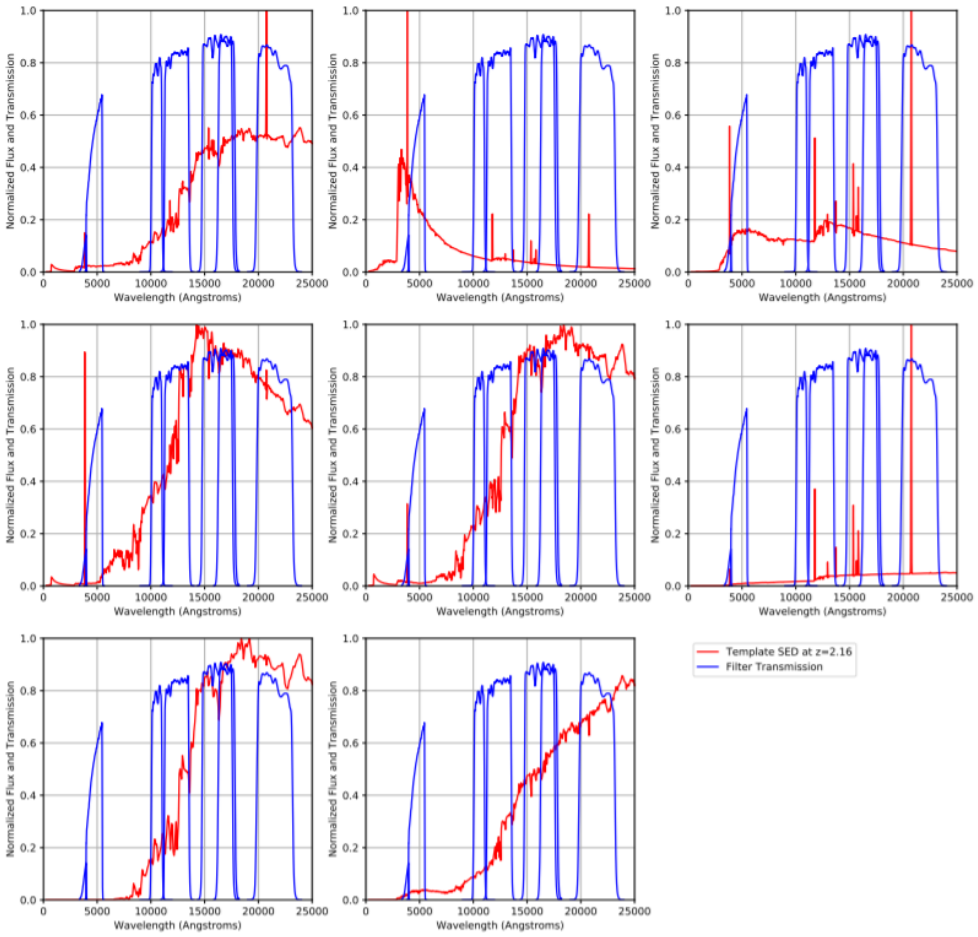


Figure 9: Template SEDs at z=2.16

Each frame shows one of the eight template SEDs used by EAZY to calculate photometric redshifts. When calculating the redshift, EAZY creates a linear combination of these eight SEDs that best fits the observed SED. Each SED shown here is shifted to how it would appear at a redshift of 2.16. The y-axis on each plot has been normalized to one so the filters show which section of the SED would appear in each filter.

EAZY then weights the chi-squared values with a prior to determine which template combination and redshift correctly models the target galaxy. A prior is a probability distribution, in this case the probability of finding a galaxy with certain properties, that describes how we expect the data to look based on previous experiments (Benítez 2000). The prior we used, shown in **Figure 10**, describes how probable it is to find a galaxy of a given brightness at a certain redshift. For example, the highest peak means that it is highly probable that the galaxy with a Ks magnitude of 19.5 is at a z~0.9. Following this line, there is almost zero chance that the galaxy with that magnitude will be found at a redshift > 2.5. We did not choose to change the prior to expect more galaxies at a redshift of 2.1, even though we know there is a higher density in this region, because we did not want to increase the rate of false positives. EAZY then chooses the template SED combination and redshift that produces the lowest chi-squared value.

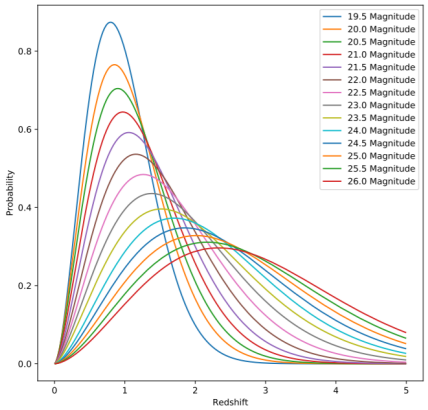


Figure 10: Ks Magnitude Prior

The Ks magnitude prior used in EAZY. The prior describes the probability of finding a galaxy at a given redshift where each line in the plot shows the probability distribution for a different Ks magnitude. Giving this prior to EAZY greatly reduces the redshift space that it must search to determine the best fit SED and reduces the number of nonphysical results.

Data Analysis

Zeropoint Comparison

Before the final catalog is used to determine photometric redshifts, we compared our results so far with previous publications. To ensure that our data is calibrated correctly, we compare total magnitudes for each filter in areas that are covered by our catalog and Tanaka 2013. The goal of this comparison is to show that the zeropoints for each filter are correct by confirming that the magnitudes of stars and bright galaxies are the same in both catalogs.

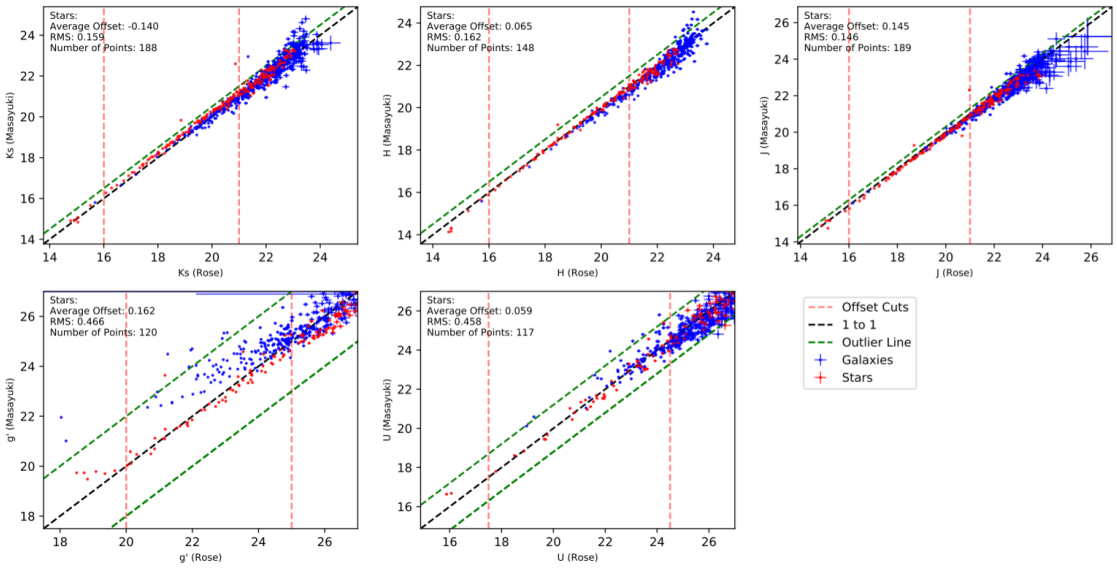


Figure 11: Zeropoint Comparison

Zeropoint comparisons for the five filters that appear in our catalog and Tanaka 2013. The dashed black line shows where objects fall when the zeropoints match. Stars, shown in red, between the dashed green and red lines are used to calculate the zeropoint offsets. A statistically significant offset was found in the J and Ks bands.

Figure 11 shows the zeropoint comparison for the five filters that are in both catalogs. The magnitudes from Dr. Demarco's catalog are estimates of total magnitudes

derived from aperture colors and a Ks total magnitude. The calculations for these total magnitudes are given by **Equation 3**. The g' filter was not converted to an estimate of the total magnitude because the data was taken on the Hubble Space Telescope. Since this data is space-based, the light from the objects is not dispersed by the atmosphere. The objects' small apparent size means the 0.5 arcsecond apertures used to measure magnitudes captures enough light to be considered a total magnitude. The u' band in our catalog was transformed to the Johnson U band, the same one in Dr. Demarco's catalog, using Equation 3 (Jester et al. 2005).

$$U = .965 * (u' - g') + g' - .78 \quad (3)$$

Outliers were checked individually to determine why their magnitudes were not consistent. Stars that are brighter than the upper-magnitude cut were removed because they were saturated in one of the two catalogs. Objects that are dimmer than the lower-magnitude cut were removed because their uncertainties became too large to accurately determine a zeropoint offset. Objects below or above the zeropoint-offset cuts were removed because their magnitude measurements were contaminated by nearby objects.

Except for J and Ks, we found no significant offset between our zeropoints and Tanaka 2013. To understand the offsets in the J and Ks filters I contacted Dr. Masayuki Tanaka, who originally created the catalog provided by Dr. Demarco. Dr. Tanaka confirmed that the zeropoints for these two filters in his catalog were not well determined and the ones from our data were more likely to be correct.

Redshifts

The observed SEDs cannot be used to uniquely determine if a galaxy has a redshift near 2.1 with the current filter measurements. By examining the galaxies that

have spectroscopically confirmed redshifts, **Figure 12**, it is hard to distinguish a galaxy at a redshift of 2.1 from a galaxy at a redshift of 1.3. At a redshift at 2.1, the H-alpha emission line falls in the Ks band, the H-beta and OIII emission lines fall in the H band, the OII emission line falls in the J band, and no emission lines fall in the Hl or J1 bands. This distribution of emission lines causes the observed SEDs to appear very similar to a galaxy at a redshift of 1.3 where the Ks band has a higher flux because a significant continuum emission is shifted into the bandpass, the H band contains the H-alpha emission line, the J band contains the H-beta and OIII emission lines, and the Hl and J1 filters do not contain any emission lines.

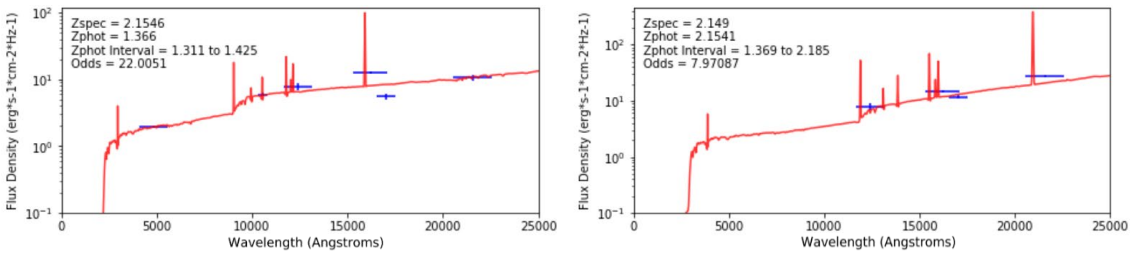
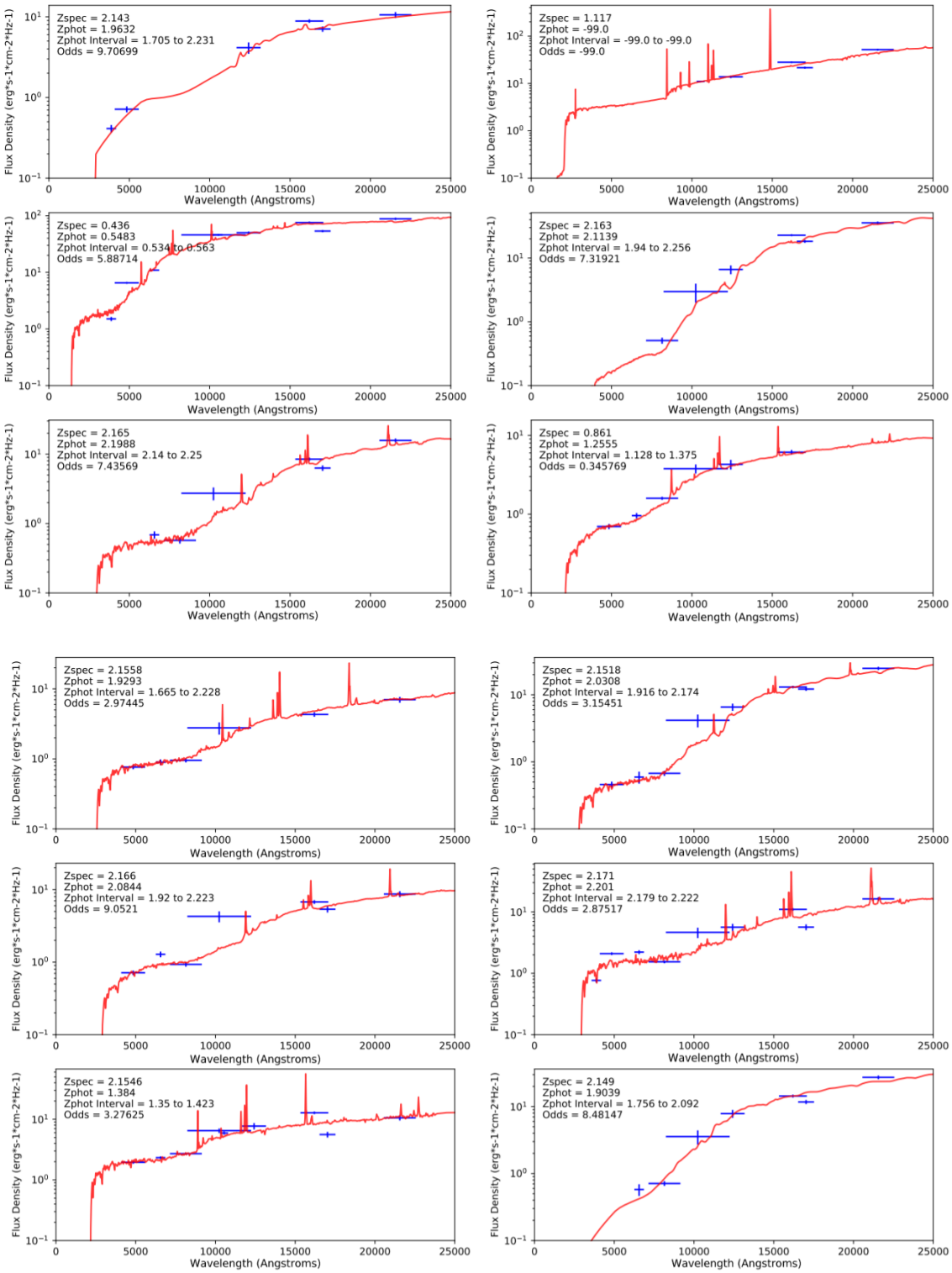


Figure 12: Redshift Degeneracy

SEDs for two galaxies both known to be at a redshift of 2.1, but with photometric redshifts at 1.3 and 2.1. The degeneracy arises from the positioning of the H-alpha, H-beta, OII, and OIII emission lines in the J, H, and Ks filters. Since this degeneracy cannot be broken with our seven filters, and changing the prior would unduly bias the results, we conclude that these seven filters cannot accurately determine whether a galaxy is located at a redshift of 2.1.

To obtain better photometric coverage, we included the filters from the catalog that Dr. Demarco provided at the beginning of the project. By including the R, I and z bands, the degeneracy between 1.3 and 2.1 redshifts was broken. **Figure 13** shows the SED fits for all galaxies that had spectroscopically confirmed redshifts. We were able to

determine the redshift of 8 of the 11 galaxies known to be at a redshift of 2.1 from their photometry.



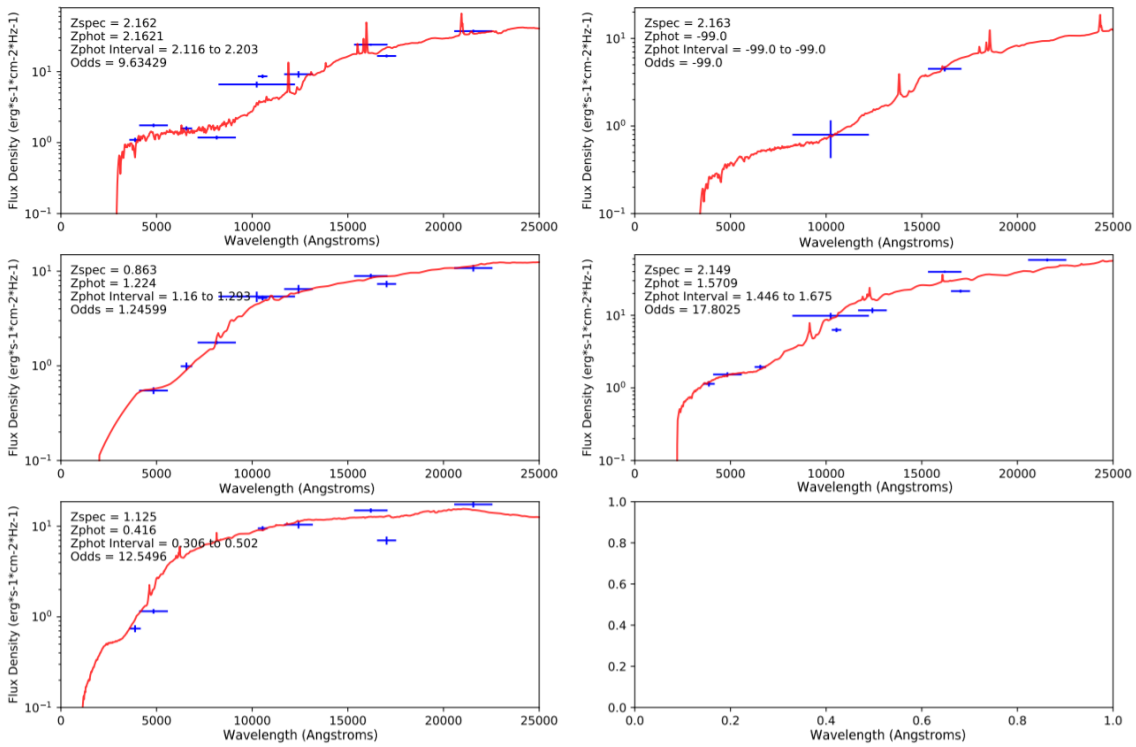


Figure 13: Spectroscopic Comparison

Template and observed SEDs for all galaxies with spectroscopically confirmed redshifts. 8 of the 11 galaxies at a redshift of 2.1 have photometric redshifts whose one-sigma confidence interval overlaps with the known redshift of the protocluster. None of the galaxies known to not have redshifts of 2.1 have photometric redshifts overlapping the redshift of the protocluster.

Based on 1-σ confidence intervals produced by EAZY for the galaxies known to be protocluster members, we defined a galaxy to be a member of the protocluster $2.1 < z < 2.18$. There does not seem to be a strong correlation between EAZY's quality factor and how well it determined the galaxy's redshift, so no bounds were placed on this criterion. Instead, we required that the object be detected in at least 6 filters to ensure that the redshift obtained was of a good quality.

The difference between the photometric and spectroscopic redshifts for the galaxies that do not have a redshift of 2.1 are higher than those of galaxies with

redshifts of 2.1. This is an expected consequence of the filters that were used to obtain the photometry. The seven original filters used were intended to be used for calculating photometric redshifts and were specifically chosen for their ability to identify galaxies at a redshift of 2.1. Accurately determining photometric redshifts for all galaxies can require more than 20 bands (Toft et al. 2009).

Completeness

To determine what magnitude our sample is complete to, we calculate limiting magnitudes for each filter. We can only use sources brighter (smaller in magnitude value) than the limiting magnitude in population statistics and characterizations because we do not otherwise know the true population. The plots used to calculate the limiting magnitude are shown in **Figure 14.** The procedure to calculate the limiting magnitude was chosen based off Zirm et al. 2008 where it is the magnitude at which the number of sources no longer increases exponentially with increasing magnitude. For a more robust value, simulated galaxy fields would have to be made and analyzed.

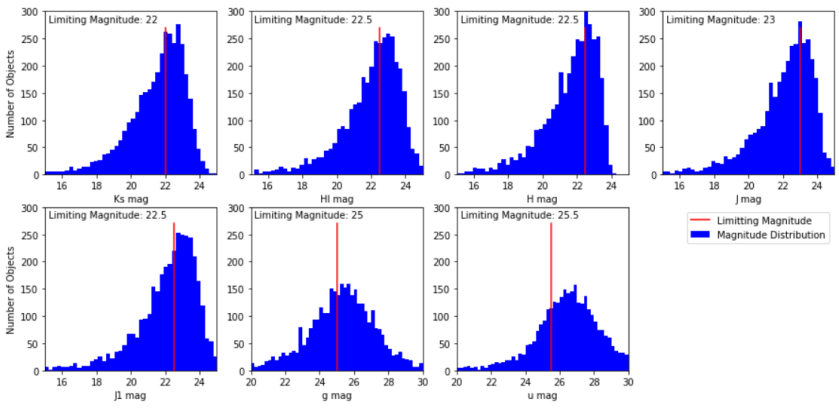


Figure 14: Limiting Magnitudes.

Magnitude distributions for all seven filters (blue histogram) with corresponding limiting magnitude (red line). The limiting magnitude is defined to be the magnitude at which the number of objects stops increasing exponentially with increasing magnitude.

Results

Of the 500 galaxies that were measured in the 10 filters used to determine photometric redshifts, 24 were determined to be at a redshift of 2.1. The redshift of 8 of these galaxies had already been determined spectroscopically, and 16 galaxies are newly determined to be at this redshift. The spatial distribution of the galaxies is shown in **Figure 15** where the Spiderweb Galaxy is at the origin. The circle shown on the plot is a calculated upper-bound on the radius of the protocluster (Kuiper et al. 2011).

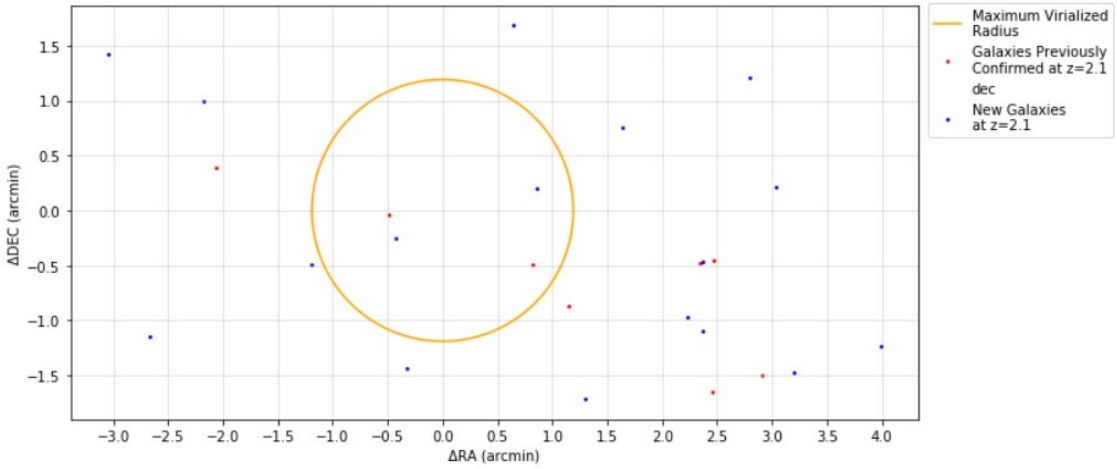


Figure 15: Galaxies With z~2.1

Spatial distribution of all galaxies whose one-sigma confidence intervals overlap with the redshift of the protocluster. Red galaxies have known spectroscopic redshifts at 2.1 and blue galaxies only have photometric redshifts. The orange circle shows the upper-bound for the space that has interacted with the core of the protocluster.

We detected two new galaxies that are possibly part of the protocluster and 14 new galaxies that are part of the field. The density of galaxies in the field, given the limiting magnitudes shown in **Figure 14**, is 1 galaxy for 3 arcmin^2.

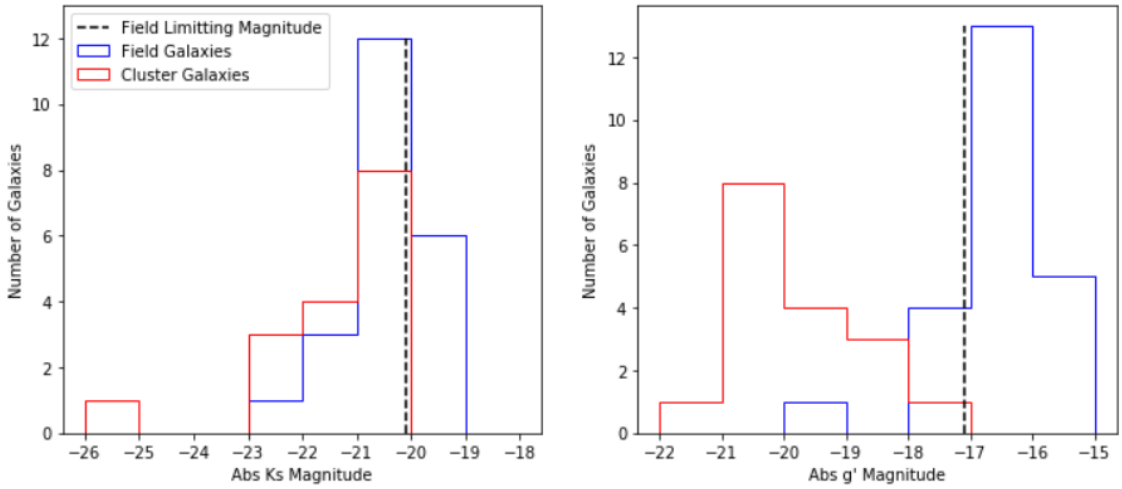


Figure 16: Cluster and Field Luminosity Distributions

Luminosity distributions for cluster (red) and field (blue) galaxies for the Ks (left) and g' (right) filters. The x-axis is the magnitude of the object if viewed from 10 parsecs, its absolute magnitude. The distributions of cluster and field galaxies in the Ks plot are statistically similar, suggesting there is no link between H-alpha emissions and being a cluster member. The distributions of cluster and field galaxies in the g' plot are statistically dissimilar, suggesting a link between Lyman-alpha emission and being a cluster member.

The luminosity functions for the Ks and g bands are shown in **Figure 12** for both the field galaxies found in this paper and the core cluster galaxies found in Hatch et al. 2009. The Ks band shows a similar distribution of galaxies between the cluster and the field, excluding the Spiderweb. This suggests that the cluster environment does not correlate as strongly with H-alpha emission than previously thought (Zirm et al. 2008). However, we acknowledge that the field sample size is likely not large enough to draw definite conclusions.

Normalizing the luminosity functions by area (not shown due to the difference in scales) shows that the core of the cluster is 300 times denser than the surrounding field. This result is 50% higher than previous estimates based on simulations (Kuiper et

al. 2011). Considering Poissonian errors, which are calculated using **Equation 4** (Hatch et al. 2009)**,** the range of

possible densities is from 240 to 400. However, the sample size is still very small and requires a larger sample from the surrounding field.

$$Counting\ Error \sim (number\ of\ galaxies\ in\ each\ bin)^{1/2} \quad (4)$$

The g-band luminosity function does not have the same distribution of galaxies in the field as it does in the core of the cluster. One reason for the different distributions is that the core of the cluster is known to have a high overdensity of Ly-alpha emitters (Kurk et al. 2004a). However, since these galaxies are above our limiting magnitude for the g band, we should have detected an overdensity of Ly-alpha emitters in the cluster. The most probable cause for not detecting them is that these galaxies are too faint in the H band to be detected above the background noise.

Discussion

Within a 4x8 arcminute rectangle around the Spiderweb Galaxy, we find 16 new galaxies with a photometric redshift of 2.1. Two of these galaxies are potentially part of the virialized mass of the cluster and 14 are likely part of the surrounding field. Following up with spectroscopic analysis of the potential cluster members will determine if this cluster is a virialized mass that formed from one overdensity or two non-virialized masses that are merging together (Kurk et al. 2004b). Following up with spectroscopic analysis of the surrounding field galaxies, alongside an analysis of the entire field region, can give a detailed picture of how the cluster formed and how it will evolve over time.

While this project has increased the sample size for both the cluster and the field, improvements can be made on the analysis of this data to provide better results. First, photometric redshifts were only able to be accurately determined in 50 percent of the field. Since the design of this project was to understand how the infalling region would affect the evolution of the cluster, this reduction in sample size greatly reduces our ability to answer this question. Second, the method used to select galaxies from the images was shown to bias against Ly-alpha emitting galaxies which are known to be abundant in the cluster.

A possible reason that we were unable to determine photometric redshifts for the entire field of view could be due to errors in the NIR zeropoints. For the J1, J, Hl, and Ks filters, there was only one standard star that was unsaturated leading to an uncertain measurement. If this is the case, the issue can be resolved by obtaining spectra for all objects in the field, convolve the spectra with the FourStar system and filters, and then

compare it to the measured magnitudes. If the data are determined to be calibrated correctly the first improvement would be to obtain significantly deeper imaging of the 11'x11' field. Additional imaging of the R, I, and z bands are needed since we showed that accurate photometric redshifts can be calculated by including measurements from these bands.

A way to reduce the bias from our selection method would be to use all master science images as detection images instead of just the H band. The best way to do this would be to run Source Extractor in dual image mode seven times, each with a different detection image to obtain 49 catalogs. Then merge the seven catalogs corresponding to the same filter and take one measurement for each of the objects in the catalog. Then combine the seven resulting catalogs into one. This method will produce more objects that are only detected in a few filters, but should decrease the sampling bias currently found in the data.

Bibliography

Benítez, N., 2000, ApJ, 536, 571

Bertin, E. & Arnouts, S., 2010, ASCL, 117, 393

Brammer, G., Dokkum, P., Coppi, P., 2008, ApJ, 686, 1503

Cousins, A., 1973, Royal Astronomical Society, 77, 223

Diehl, T., 2012, Physics Procedia, 37, 1332

Erben, T., Schirmer, M., Dietrich, J. et al. 2005, AN, 326, 432

Gilmore, G., Randich, S., Asplund, M., et al., 2012, The Messenger, 147, 25

Hatch, N., Overzier, R., Kurk, J., et al., 2009, RAS, 395, 114

Jester, S., Schneider, D., Richards, G., 2005, ApJ, 130, 873

Jørgensen, I., Chiboucas, K., Nielsen, L., Takamiya, M., 2018, ApJ, 235, 29

Kodama, T., Tanaka, M., Tanaka, I., Kajisawa, M., 2007, Cosmic Frontiers, 379, 202

Kuiper, E., Hatch, N., Miley, G., et al., 2011, Royal Astronomical Society, 415, 245

Kurk, J., Rottgering, H., Pentericci, L., et al., 2003, New Astronomy Reviews, 47, 4, 339

Kurk, J., Pentericci, L., Overzier., L. et al., 2004b, A&A, 428, 817

Kurk, J., Pentericci, L., Rottgering, H., Miley, G., 2004a, A&A, 428, 793

Li, T., DePoly D., Marshall, J., Nagasawa, D., et al., 2014, SPIE Astronomical Telescopes + Instrumentation, 9147

Miley, G., Overzier, R., Zirm, A., et al., 2006, ApJ, 650, 29

Oke, J., 1964, ApJ, 140, 689

Padmanabhan, N., Saljek, U., Strauss, M., et al., 2004, *New Astronomy*, 9, 329

Persson, S., Murphy, D., Birk, C., et al., 2013, ApJ, 125, 654

Persson, S., Murphy, D., Gunnels, S., et al., 2002, ApJ, 124, 619

Schirmer M. 2013, ApJS, 209, 21

Tanaka, M., Toft, S., Marchesini, D., et al., 2013, ApJ, 772, 113

Toft, S., Franx, M., van Dokkum, P., et al., 2009, ApJ, 705, 255

Wells, D. & Greisen, W., 1979, International Workshop on Image Processing in Astronomy, 64, 445

Zirm, A., Stanford, S., Postman, M., et al., 2008, ApJ, 680, 224

Zirm, A., Toft, S., Tanaka, M., 2012, ApJ, 744, 181